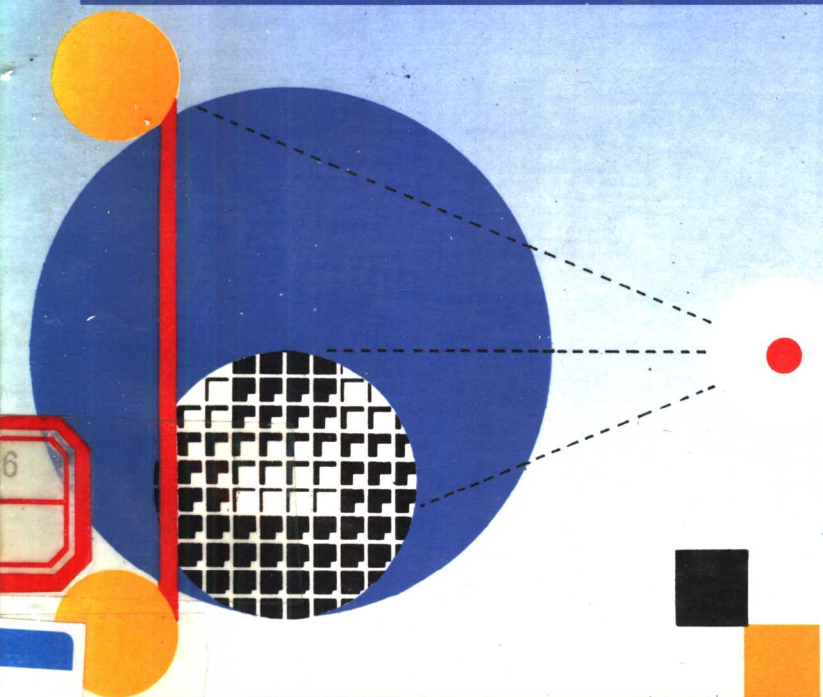


晨旭编著

中学数学考试命题研究

Zhongxue Shuxue
Kaoshi Mingti Yanjiu



湖南教育出版社

G633.6/471

中学数学考试命题研究

Zhongxue Shuxue
Kaoshi Mingti Yanjiu

晨旭编著

湖南教育出版社

中学数学考试命题研究

晨 旭 编著

责任编辑：郑绍辉

湖南教育出版社出版发行

湖南省新华书店经销 湖南省新华印刷二厂印刷

787×1092 毫米 32 开 印张：9.625 字数：229,000

1997 年 5 月第 1 版 1997 年 5 月第 1 次印刷

ISBN7-5355-2461-3/G·2456

定价：9.40 元

本书若有印刷、装订错误，可向承印厂调换

前 言

考试作为一种检测手段，在学生学习成绩的考查和人才的甄别、选拔中被广泛使用。考试不仅是学校中一项不可缺少的经常性的工作，而且在社会生活中也存在着各种大规模的入学统一考试、资格考试和水平考试。各类考试的组织和实施，以及考试结果的使用，为许多人所关注和研究。

长期以来，我们在中学数学考试命题的实践，及其理论研究上，做过一些工作，深感中学数学命题是一项科学性、技术性都很强的工作。为了做好这项工作，不仅在数学方面必须具备坚实的基础和丰富的教学经验，而且对教育测量学、教育统计学和考试学也应有较深入的了解，并善于结合数学学科的特点，加以运用。本书作为我们的研究心得，着重讨论下述问题：

一、从考试命题这个层面出发，概括地讨论考试中的一些基本问题，包括考试的性质和分类；考试命题的意义和原则；考试大纲的制订和使用；分数的转化和使用；考试质量评价。

二、较为深入和系统地探讨了中学数学考试命题的理论和实践，包括中学数学试题的分类；试题和试卷的编制方法及其技巧；命题中常见的失误及其防止办法；试题和试卷的难度估计与预测。

三、试题和试卷使用后的质量分析，包括定性和定量的分析方法及其技术。

本书作者有：任子朝、陈云烽、陈大钧、毛经中和钱昌本。我们恳请读者批评指正，提出宝贵意见。

作者

一九九六年十二月

目 录

第一章 考试的类型与目标	(1)
1.1 考试的一般分类	(1)
1.2 目标参照性考试	(6)
1.3 常模参照性考试	(10)
1.4 中学阶段的数学考试	(13)
第二章 考试命题的意义与原则	(26)
2.1 考试命题的意义	(26)
2.2 考试大纲	(30)
2.3 考试命题的原则	(34)
2.4 命题的依据——双向细目表	(43)
2.5 中学数学考试的命题	(51)
第三章 试题	(66)
3.1 试题功能分析	(66)
3.2 试题结构特点	(74)
3.3 试题分类方法	(83)
3.4 试题质量要求	(104)
3.5 试题编制规律	(111)
3.6 试题的难度预测	(124)
第四章 试卷	(130)
4.1 试卷编制的一般原则	(131)
4.2 试卷的设计	(139)

4.3	目标性考试试卷和常模性考试试卷	(145)
4.4	试卷的编制	(148)
4.5	试卷编制工作中的几个问题	(161)
第五章	命题的方法与技巧.....	(177)
5.1	命题的目的	(177)
5.2	选题	(180)
5.3	改题	(190)
5.4	编题	(214)
5.5	关于编制试题的几个问题	(246)
第六章	考试的质量分析.....	(253)
6.1	考试的质量分析的意义和作用	(253)
6.2	特征量数	(257)
6.3	试题的质量分析	(269)
6.4	试卷的质量分析	(278)

第一章 考试的类型与目标

1.1 考试的一般分类

试题用于考试，不同的考试对试题有不同的要求。因此，在着手编拟和命制试题之前，必须明确使用该试题的考试之性质，了解其是什么类型的考试。

考试的类型，可按不同的角度进行区分。例如，从用途上区分，一般可分为成绩考试、水平考试、学能考试、诊断性考试等；从分数标定的方法上区分，可分为目标参照性考试和常模参照性考试；从考试方法上区分，可分为笔试、口试和操作考试；从考试方式上区分，可分为综合式考试和分离式考试；从考试规模和考试时间上区分，可分为课堂小测验、学习中的阶段性考试，大规模的会考、统考或竞赛考试；从考试要求上区分，可分为能力考试、速度考试；等等。

在命题这个层面上看考试的类型，最为关心的是考试的用途、要求、性质和目标，它决定了命题的内容及其难度控制；其次是考试的方式、方法、手段和规模，它制约着试题的题型、题量和结构；再次，是反映考生成绩的分数制度，它影响着试卷的结构、长短、难度分布，以及测试误差的调控。

与中学数学命题有关的考试，从考试的用途、要求和目标上区分，必须注意下列各类考试的特点和区别，以便更好地指导命题。

1.1.1 诊断性考试

这类考试，是在测量学生掌握某一部分教学内容的使用，是一种具有诊断性质的考试。通过这种考试可及时看到学生的学习情况，直接获得教学的反馈信息，调整教学的进度和教学的方式、方法。这种考试一般在课堂上进行，学生的情况，通常比较相近。考试的安排，多为主讲教师根据教学和学生情况自行作出，试题也多由教师根据本人教学的需要自行命题，灵活性比较大。学生的成绩，主要作为了解教学情况之用，不一定作为衡量水平之用。这种考试的范围比较小，内容比较集中，目标单一。试题往往只围绕一个核心进行命题，关键是能否抓住考查内容的重点和难点，抓住了才能促进教学。在试题的设计上，要十分注意其诊断作用，即：通过试题的解答，能清楚地反映学生对某一特定的知识点和数学方法是否理解和掌握，有时还得反映其熟练程度和深透程度。一般说来，不宜采用综合性试题。例如，为了检查学生掌握对数函数性质的情况，与其使用试题：

$$\text{解不等式 } 2 + \log_2(5-x) + \log_2 \frac{1}{x} > 0.$$

倒不如改用下面的一组试题：

- (1) 已知 $a = \log_2 3 + \log_2 5$ ，求 2^a 的值；
- (2) 设 $\log_2 y = \log_2(5-x)$ ，求用 x 表示 y 的函数解析式（要求不含对数符号）；
- (3) 解方程 $2\log_3 x + \log_3(x^2 - 2x + 4) = 0$ ；
- (4) 求实数 a 的取值范围，使方程 $2\log_3 x + \log_3(x^2 - ax + 4) = 0$ 有解；
- (5) 设 $a > 0$ 且 $a \neq 1$ ，比较 $\log_a 2$ 与 $\log_a 3$ 的大小。

这是因为后面的这组试题，能较好地起到诊断作用，它的每道试题都比较集中地考查对数函数的各项性质。而前面解对数不等式的那道试题，则综合地考查了对数换底公式、乘除的对数公式、方幂的对数公式、对数的单调性，以及解分式（有理函数）不等式等多个知识点，及其灵活运用能力。因此，只有在高一层次上进行诊断时，它才起作用。也就是说，如果在确知学生大体上已掌握了对数函数的性质，为了检查他们是否能够灵活运用时，使用该题才较为恰当。

1.1.2 成绩考试

学生在一定阶段的时间内，完成了某一教学大纲或教材的学习，为了检测他们的学习成绩，往往对其进行成绩考试。例如，每学期的期中考、期末考，甚至升级、毕业考试，都属于这类考试。这是一种检查学习进度的考试，考试内容为所考查的阶段内的教学内容，试题的命制，可看作对所考核的教学内容进行抽样。这种抽样，不宜认为是随机抽样，而应该认为是：能比较好地反映教学内容的全貌的抽样。因此，必须对教学内容和教学要求作深入的分析，弄清其中的主次轻重，命题时应把握住基本的和重点的内容，而且在设问上也应考虑不同的层次要求，使之通过考试能较好的反映学生对所学内容的掌握程度，包括学到手的知识之多寡，以及理解和掌握的程度。如果采用百分制来评定成绩，得60分也就应该是相当于掌握了60%左右的教学内容。如果试题的知识复盖面太少，或者复盖的多为次要部分，则难以实现这样的考试目的。在试题的难度控制上也有同样的问题，太浅、太深都会使考试失效。

这类考试，通常是按班级进行，而且，多数情况下，由教师根据所教的内容自行命题。因此，校际之间的考试成绩，往

往难以比较，也难以用它来评估学生的水平，因为它缺乏统一的标准。

1.1.3 水平考试

旨在测量考生是否达到某一水平的考试称为水平考试。它有明确考查标准，在较大的时空范围内，用统一的标准，组织和实施考试。目前，在我国实行的普通高级中学的毕业会考，就是一种国家承认的省级水平考试。

水平考试的命题，既要考虑到考生学过（或应该学过）一些什么内容，但又不能局限于某一特定教材作为依据，它必须以考试大纲规定的统一标准作为依据。这样，考试才不会偏颇于某些学校或单位，考试的成绩才具可比性。只要达到了规定的统一标准，不同单位或考生可各显神通，采用不同的教材和教学方法。

这类考试的试题，应力求较好地反映考试所规定的水平标准，使之通过考试，能有效地确定每个考生是否达标，达标者确能适应新的学习或工作，不达标者也能明白自己的不足，及其努力的方向。因此，在命题之前，必须深入研究所规定的统一标准，探讨怎样的试题和试卷才能反映规定的标准，试题和试卷是标准的具体反映，从知识点的采样，到能力的要求，都要很好地把握分寸。

这类考试的分数制度，通常可用等级制，关键是达标和未达标两大级，在两大级之中，再适当分成若干级别，一般不必太细。

1.1.4 学能考试

这类考试，旨在测量考生完成某项任务的能力倾向，这些

能力倾向，在很大的程度上带有潜在性，即在考试时，考生也不一定已具备这些能力，而只是具备发展这些能力的基础和倾向。

这类考试的命题并不完全根据以往的教学内容来命题，更主要的依据是对学能结构的分析，对以往的教学内容，作为命题的取材，往往是选用那些可作为学能基础的内容，在要求上也不一定局限于教学要求这个层次上。

这类考试往往带有选拔的性质，应有较好的区分度，使能力倾向的大小和优劣能得到较为细致的区分。

各类入学考试，其实质都可作为学能考试来对待。水平考试与学能考试的基本区别在于：前者着重的是现时所达到的水平，后者所看重的是未来的发展倾向和趋势。从这一角度看，我国普通高等学校入学统一考试（简称高考），由于历史的原因，在高中毕业会考制度建立之前，它既有水平考试的特征，又兼有学能考试的选拔性质。随着高中毕业会考制度的实施和成熟，高考学能考试的性质日显突出，能力和学习潜力的考查，在高考中，正在得到加强。

1.1.5 速度考试

在各类考试中，对解题速度都有一定的要求，因为任何考试都得在规定的时间内解答一定数量的试题，这本身就表明了速度的要求。所谓速度考试，指的是测量考生完成某一任务的速度之考试。速度的要求往往是在能力要求的基础上提出的，突出速度的考试，在能力要求上往往比较集中单一。例如，打字速度的考试，数和式的四则运算的速度考试，解方程的速度考试，作图速度的考试，等等。对于大规模的综合性考试，一般较难把速度要求放在首位。

速度考试的试题，其题量往往比较大，对速度也往往提出定量化要求。例如，打字速度考试，往往明确一个小时要完成多少个字符的任务；计算速度考试，通常也会规定一小时要完成多少个算式的计算任务。

速度的快慢往往反映了能力的强弱，因此，在一些大规模的考试中，往往含有较多的考查项目和试题，其用意不在于要求考生都能全部完成，主要的目的是借助它测量考生的解答速度，从一个重要的侧面考核考生的能力，提高考试的区分度。这一点，对各类考试的命题，有着普遍的指导意义。

1.1.6 竞赛考试

也可叫做竞技考试，这类考试通常用于特定群体的专项能力和技能（或者某类综合能力和技能）的测试。测试的目的，在于把最优秀的那一部分人突现出来，是竞争性最强的一类考试。

这类考试的命题，自由度比较大，并且往往以高难度的试题为主。当然，它也仍然要从参加考试的群体之实际出发。因为如果试题难到无人能够解答，那么试题也如同虚设。命题的依据是竞赛的目的和要求，而不是教材和教学。

在我国，与高中数学教育相关的规模最大的两项考试是高中毕业会考和高考，可分别归属于目标参照性考试和常模参照性考试。下面两节，将分别讨论这两类考试的特点、性质和命题要求。

1.2 目标参照性考试

目标参照性考试，也有人称之为目标参考性考试，或者尺度参考性考试。

凡是评分时参照某一事先规定好的尺度标准和目标，并用评出的分数作为考生成绩的反映和标识，这样的考试统称为目标参照性考试。

一般说来，单科（或单项）的成绩考试和水平考试多属目标参照性考试，而多科的水平考试，则往往介乎于目标参照性考试与常模参照性考试之间。

目标参照性考试所给的分数，并不考虑其他考生的情况，只是相对于给定的标准加以评定。只有达到规定的目标标准或尺度标准才算通过，或称及格、达标。分数的解释离不开具体的试题，孤立的一个分数也反映不出考生成绩在全体考生中所处的地位。不同学校、不同时间的考试成绩，缺乏可比性。

在目标参照性考试中，可能全体考生都达到及格，也可能全部不及格。社会上许多任职资格考试，也都属于目标参照性考试。通过考试，并达到及格，才能领取从事该项工作的证书或执照。

1.2.1 目标参照性考试的组织和实施

以班级为单位，结合教学进度所进行的成绩考试，这类小规模的目标参照性考试的组织和实施，一般由教师或学校，结合学生和教学的实际，负责安排和开展，其依据主要是教学大纲规定的教学要求。

对于大范围和大规模的目标参照性考试，一般都应有专门的机构负责组织和施行。例如，我国实行的高中毕业会考，便是由各省（自治区、直辖市）的教育行政部门成立专门班子，负责组织和实施。各省对这班子的称谓也不一致，有叫会考办的，也有叫考试院或考试中心的。不过，具体的任务和工作环节大体相同，主要是：考试大纲的制定和公布；试题的征集和试卷

的制定；试场的安排和考试的施行；评卷工作的组织和进行；考生成绩的公布；补考的安排和实施。

为了保证考试的权威性，上述的每个考试环节都不能掉以轻心，都必须从严要求。但作为目标参照性考试，目标或尺度标准的制定，是至关重要的，它不仅影响着考试的质量及成败，而且对社会生活也存在着重大的影响。目标和尺度的规定如果科学、合理、切合实际，则有利于国家教育事业的发展，对社会的进步起促进的作用，如果欠科学、不合理、脱离实际，则会起着相反的作用，甚至可能引发社会问题。以高中毕业会考为例来说，既然它是省级的水平性考试，因此，尺度标准的制定，就必须结合本省的高中教育的实际，既不应把标准定得太高，但又不能迁就实际中的落后面。还得以国家教委所颁布的教学大纲的教学要求为依据，使得只要具备基本的办学条件，经过正常的教学，达到基本教学要求的考生，都能达到及格的要求。这样才能促进高中教育的发展，维护社会的稳定。可见，各省的高中毕业会考的尺度标准，并非完全一致，而是在大体相同的前提下，允许在一定范围内存在差别。

尺度标准的制定是一项十分细致的工作，也是一项专业性和技术性都很强的工作。没有深入广泛的调查研究，没有一定的人力、物力、财力和时间的投入和积累，断然是无法做好的。

1.2.2 目标参照性考试的命题

目标参照性考试的尺度标准的规定，通常是原则规定，为了说明这些原则规定，往往辅以试题样本，有的是逐题举例说明，有的是整卷展现说明。但所有样题，都不是考试时正式出现的试题，这是由于任何考试的试题在考前都是保密的缘故。就是按同一标准，在不同时间所进行的考试，其试题也不一样。可

见,在同一尺度标准规定下,可以命制出许多份不同的试卷.命题时,应力求使每份试卷尽可能接近规定的标准.这里,之所以采用“尽可能接近”的提法,是因为,要准确地与规定的标准保持一致,根本无法做到.考试的尺度标准,不同于物理量的尺度标准,其可变因素和弹性要大得多,加之作为考试测量量具的试题,往往只能使用一次,使用前又得严格保密,而对使用效果的了解,又得在考试之后,试题已经“报废”之时,才能进行.这一切,使得“试题要准确反映规定的尺度标准”的提法不切实际.这一点,在命题时,应清醒地意识到,以便促使命题者更为深入地做好调查研究工作,对规定的尺度标准逐条逐项加以理解和把握,适当细化,分清主次和轻重,同时十分注意命题技术的讲究和运用,借以提高命题的质量.

同一项目标参照性考试,往往不止进行一次,而是周期性定期或者不定期地进行多次,因而对所采用的试题应力求可比,保持稳定.这也是这类考试对命题的一个基本要求.

1.2.3 目标参照性考试的分数制度

目标参照性考试的分数制度一般采用百分制或等级制(有二级制、四级制、五级制等等).

在百分制中,通常以60分为及格线.但在实际运用中,尤其是大规模的考试,往往会出现试题偏易或偏难的情况,因此,调整及格线的做法,也是常有的事.如果调整的幅度很大,则说明命题的质量较差,必须及时总结经验教训,改进命题工作.

二级制,通常分为及格(或合格)与不及格(或不合格)两级;四级制,通常是在及格这一级中细分为优秀、良好和及格三级;在四级制的基础上,又把不及格段细分为二级,便成为五级的记分制.

记分制粗细的采用，往往取决于考试的用途。由于目标参照性考试，分数的评定是参照预先规定的标准进行，因此，不论采用的是何种分数制度，都应按标准给分，不应随意改变标准，不应随人而异，才能保证考试的公正。

1.3 常模参照性考试

常模参照性考试，有人也叫常模参考性考试。在这种考试中，反映考生成绩的分数与目标参照性考试不同，它不是以事先制定的尺度标准为参照点，而是参照某个常模来反映考生的分数。这里所说的常模是指：某一考生群在考试中的成绩（通常，用该群考生的卷面分数的平均分与标准差来刻划这个成绩）。也就是说，常模参照性考试中，反映考生成绩的分数是结合其它考生的成绩来给出的，从所给的分数便可看出考生在某群考生中的地位，是优是劣，一目了然。这样做，对于以选拔为目的的大规模的考试，无疑是必要的。这是因为：选拔一个考生，就必须将他与其他的考生进行比较，分辨优劣，才能决定取舍，不能光看他一个人的考试卷面分数。这个道理显而易见。事实上，在常模参照性考试这个概念及其给分方法提出之前，人们已经明白这个道理，也已懂得用排名次的方法解决这个问题。即是说，用序数来反映考生的考试成绩。这种做法，当考生的群体比较小，人数不太多时，如一个班或者一个年级，比较可行。但是，当群体很大，考生的人数成千上万时，如一个市、一个省，就难办了。这时，得同一卷面分数的考生往往不止一个，甚至可以是数百人，要对他们进行排序，根本不可能，而且，这时的排序，往往也失却意义。用序数表示成绩，也同样弄不清他在全体考生中的位置。因此，给其序数，倒不如给

出考生在群体中的百分位，反倒说明问题。但是，用百分位，实行起来也比较麻烦。这便引发了用分数反映百分位的做法，产生了“常模”和“标准分”的概念。对此，在下面讨论分数制度时，再作进一步的介绍。

1.3.1 常模参照性考试的组织和实施

常模参照性考试，一般为大规模的选拔性考试所采用，有时，也为某些大规模的水平性考试所采用。因此，必须由专门的机构负责组织和实施。其考试环节，大体上与大规模的目标参照性考试的考试环节一样。所不同的是，在完成试卷评阅之后，不是直接公布卷面的分数作为考生的成绩，而必须根据所建立的常模，把分数加以转换，变成“标准分”之后，再行公布。这一环节的实施，没有计算机和计算技术的支持，那是无法实现的。因此，可以说，常模参照性考试是现代教育测量的产物。

1.3.2 常模参照性考试的命题

作为命题的原则和要求，常模参照性考试与目标参照性考试并无根本的区别。在常模参照性考试中，考试的内容和要求，同样必须根据考试的目的和用途，给予明确的规定。命题时，同样要力求准确和稳定地反映这种规定，试题不能忽易忽难，波动太大。同时，命题时，必须更好地了解考生的实际，提高试题的区分度，使考试结果的卷面分数（通常称之为原始分），能尽可能切实地反映考生的成绩，这是建立“常模”的基础。

根据数理统计学的原理，当参加考试的考生成千上万时，其考生的成绩分布必然是正态分布，如果满分是100分的考试，平均分应为50分左右，标准差应为17分左右。如果平均分偏离